

# Achievable Rates of Multidimensional Rotationally Invariant Distributions

Johnny Karout, *Senior Member, IEEE*, René-Jean Essiambre, *Fellow, IEEE*, Erik Agrell, *Senior Member, IEEE*, and Antonia Tulino, *Fellow, IEEE*

**Abstract**—The maximum achievable rate or mutual information of multidimensional rotationally invariant distributions in the presence of additive white Gaussian noise is analyzed. A simple expression for the special case of multisphere distributions is derived. Such distributions consist of points in a multidimensional Euclidean space that are uniformly distributed over several multidimensional concentric hyperspheres. For the 2-dimensional case, such distributions have been previously considered in order to reduce the computational complexity of finding a bound on the channel capacity of fiber-optic channels. These distributions take advantage of the statistical rotational invariance of the noise and nonlinear distortions in fiber-optic channels. Using the derived mutual information expression, 2- and 4-dimensional multisphere distributions are compared for fiber-optic dual-polarization channels dominated by linear noise. At high signal-to-noise ratios, 4-dimensional multisphere distributions offer higher achievable rates than 2-dimensional ones transmitted on each of the two polarizations of the optical carrier. Such 4-dimensional multisphere distributions are also shown to be statistically invariant under 4-dimensional nonlinear transmission in fibers.

**Index Terms**—Coherent detection, fiber-optic communication, multidimensional constellation, mutual information, polarization, ring constellation, spherical distribution.

## I. INTRODUCTION

THE CHOICE of modulation formats plays a key role in meeting the demands of future communication systems. Typically, the performance of a modulation format or its equivalent constellation representation in signal space is measured by its spectral efficiency, power efficiency, and complexity. The underlying challenge is to design a constellation that provides the appropriate trade-off between these key performance measures. This draws the attention to the use of multidimensional ( $N$ -D) constellations, which provide good spectral and power efficiency at the expense of increased complexity. The dimensions, which can be realized by exploiting diversity in time, frequency (wavelength), space (fiber cores, fiber modes,

or antennas), and/or polarizations, can be looked upon as degrees of freedom used in designing constellations in their respective signal spaces.

For channels where the noise has rotationally invariant statistics, it is attractive, at least theoretically, to study input distributions that are also rotationally invariant. The multidimensional Gaussian distribution is the most common such distribution, and it is capacity-achieving over the additive white Gaussian noise (AWGN) channel under an average power constraint [1, Sec. 24], [2, Ch. 9]. Under other circumstances, it may be better to apply input distributions that are discrete in amplitude, but still rotationally invariant, i.e., continuous and uniform in phase (where “phase” may be interpreted in a multidimensional sense). Such distributions, which are called *multisphere distributions* or in the 2-D case *multiring distributions* and exemplified in Fig. 1, have received considerable interest in the literature and are further analyzed in this paper.

Shamai and Bar-David proved in 1995 [3] that multiring distributions are capacity-achieving over the 2-D AWGN channel under simultaneous average and peak power constraints, thereby generalizing a classical 1-D result by Smith [4]. Katz and Shamai [5] proved that multiring distributions achieve the capacity of the partially coherent AWGN channel. Chan *et al.* [6] considered a wide range of channels with peak input constraints and identified several cases where multiring or multisphere distributions are capacity-achieving, including two cases of the Rayleigh-fading channel, the Gaussian interference channel, and a set of parallel AWGN channels with an average total power constraint. Gursoy *et al.* [7] proved similar results for the Rician fading channel under fourth-moment or peak-power input constraints. See references in, e.g., [5] and [6] for further examples of discrete-amplitude capacity-achieving distributions.

Another application of multiring distributions arises in fiber-optic communication systems with wavelength-division multiplexing, in which the signals on different wavelengths interfere with each other nonlinearly. It was shown by Ghazlan and Kramer [8], [9] that by choosing all input distributions as multiring constellations, and under idealized channel conditions, the cross-channel interference can be eliminated.

In [10], [11], Essiambre *et al.* used multiring distributions to derive lower bounds on the capacity of optical fibers. The analysis is simplified by the fact that all points on the same ring are statistically equivalent. Therefore, a single probability distribution function can be calculated for each ring by accumulating statistics using all points on that ring.

This work was supported by the Swedish Foundation for Strategic Research under Grant RE07-0026, the Ericsson Research Foundation, the Friends of Chalmers Foundation, and the Royal Society of Arts and Sciences in Gothenburg.

J. Karout was with Bell Labs, Nokia, Holmdel, NJ 07733 USA, and Chalmers University of Technology, SE-412 96 Gothenburg, Sweden, when this work was done. He is now with Ericsson Research, Ericsson AB, 164 80 Stockholm, Sweden (e-mail: johnny.karout@ericsson.com).

E. Agrell is with the Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: agrell@chalmers.se).

R.-J. Essiambre and A. Tulino are with Bell Labs, Nokia, Holmdel, NJ 07733 USA (e-mail: {rene.essiambre, a.tulino}@nokia.com).

The mutual information was numerically computed in [10] for a nonlinear fiber-optic channel model in an optically-routed network. The study was expanded in [11], where results for the AWGN channel were included. Multiring constellations have also been used to numerically estimate the capacity of the zero-dispersion fiber-optic channel [12] and to lower-bound the capacity of the finite-memory Gaussian noise channel [13].

Coherent optical systems have a 4-D signal space, corresponding to the two quadratures in both polarizations of an optical signal. This implies that a 2-D multisphere distribution can be sent on both polarizations simultaneously, thus increasing spectral efficiency. A first question that arises is whether a 4-D multisphere distribution provides higher achievable rates for the same power efficiency compared to a 2-D multisphere distribution transmitted on each of the two polarizations. A second question is whether points on each sphere in a 4-D multisphere distribution are statistically equivalent for nonlinear propagation in fibers.

In this paper, we analytically derive the mutual information using an  $N$ -D multisphere distribution for the AWGN channel and display the mutual information results for 2-D and 4-D multisphere distributions. This mutual information serves as a lower bound on the multisphere-constrained capacity, since there is no maximization involved over the radii and the probability of each hypersphere. Finally, we show the statistical invariance of the stochastic Manakov equations that describe nonlinear propagation over optical fibers under arbitrary 4-D rotations. This shows that 4-D hyperspheres are particularly important to consider for fiber-optic communication systems.

## II. SYSTEM MODEL AND MUTUAL INFORMATION

We consider a discrete-time AWGN channel

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \quad (1)$$

where  $\mathbf{X}$  is an  $N$ -D real input vector and  $\mathbf{N}$  is an  $N$ -D normally distributed noise vector with mean 0 and variance  $\sigma^2 = N_0/2$  per real dimension, where  $N_0/2$  is the double-sided power spectral density of the underlying continuous-time noise process.

The mutual information between input and output is [2, Sec. 8.4–8.5]

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}), \quad (2)$$

where the differential entropies are

$$h(\mathbf{Y}) = - \int f_{\mathbf{Y}}(\mathbf{y}) \log_2 f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}, \quad (3)$$

$$h(\mathbf{Y}|\mathbf{X}) = h(\mathbf{N}) = \frac{N}{2} \log_2 2\pi e \sigma^2. \quad (4)$$

The maximum mutual information under an average power constraint is the AWGN channel capacity [14, Eq. (11)]

$$C = \frac{N}{2} \log_2 \left( 1 + \frac{2}{N} A \right), \quad (5)$$

where  $A = \mathbb{E}[\|\mathbf{X}\|^2]/N_0$  denotes the signal-to-noise ratio (SNR).

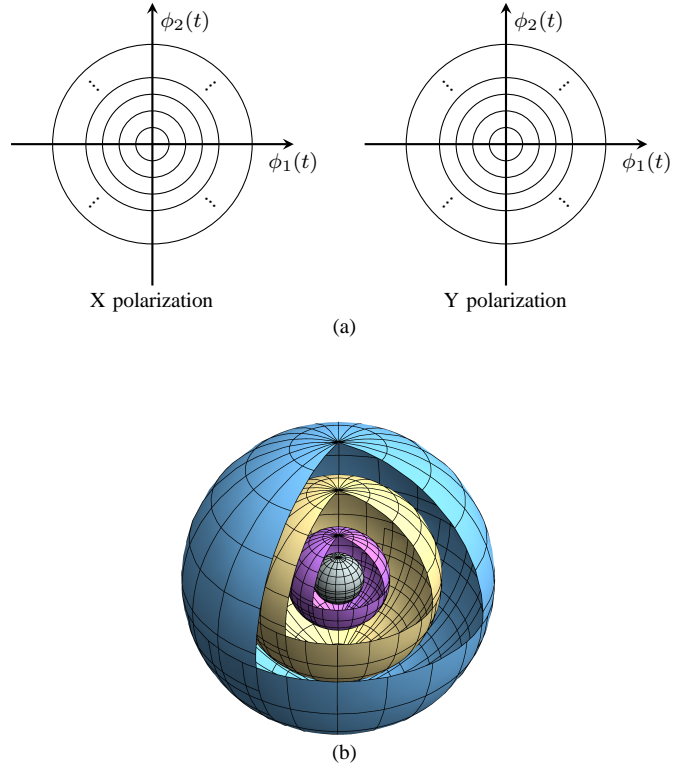


Fig. 1. (a) A 2-D multisphere distribution (multiring distribution) in both polarizations, where  $\phi_1(t)$  and  $\phi_2(t)$  are the basis functions that span the signal space of conventional in-phase and quadrature modulation formats. (b) A 3-D multisphere distribution, or a cross-section of an  $N$ -D multisphere distribution.

## III. ROTATIONALLY INVARIANT DISTRIBUTIONS

We are interested in the case when the input vector  $\mathbf{X}$  is statistically invariant under  $N$ -D rotations. Hence, its probability density function (pdf) is fully characterized by the pdf of the scalars  $\|\mathbf{X}\|$  or  $\|\mathbf{X}\|^2$ . Since the Gaussian vector  $\mathbf{N}$  is also rotationally invariant, so is  $\mathbf{Y}$ .

The following lemma characterizes the differential entropy of rotationally invariant random variables, generalizing a lemma by Lapidot and Moser [15, Eqs. (319)–(320)] to arbitrary dimensions  $N$ .

*Lemma 1:* If  $\mathbf{Y}$  is a random  $N$ -D real vector with rotationally invariant statistics, then its differential entropy is<sup>1</sup>

$$h(\mathbf{Y}) = h(\|\mathbf{Y}\|) + (N-1)\mathbb{E}[\log_2 \|\mathbf{Y}\|] + \log_2 \frac{2\pi^{N/2}}{\Gamma(N/2)} \quad (6)$$

$$= h(\|\mathbf{Y}\|^N) + \log_2 \frac{\pi^{N/2}}{\Gamma(N/2+1)}. \quad (7)$$

*Proof:* For any  $N$ -D real random vector  $\mathbf{Y}$ , the distribution of its length  $R = \|\mathbf{Y}\|$  is calculated by integrating  $f_{\mathbf{Y}}(\mathbf{y})$

<sup>1</sup>The gamma function satisfies

$$\Gamma(N/2) = \begin{cases} (N/2-1)!, & \text{if } N \text{ is even,} \\ \frac{\sqrt{\pi}(N-1)!}{2^{N-1}(\frac{N-1}{2})!}, & \text{if } N \text{ is odd.} \end{cases}$$

over all  $\mathbf{y}$  with the same length, i.e.,

$$f_R(r) = \int_{P_r} f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}, \quad (8)$$

where  $P_r$  is the  $N$ -D hypersphere with radius  $r$  centered at the origin. In the special case when  $\mathbf{Y}$  is rotationally invariant,  $f_{\mathbf{Y}}(\mathbf{y})$  is constant for all points  $\mathbf{y}$  with the same radius  $\|\mathbf{y}\| = r$ . For such distributions,  $f_{\mathbf{Y}}(\mathbf{y})$  in (8) does not depend on  $\mathbf{y}$  and can be moved outside the integral. Thus, for any  $\mathbf{y}$ ,

$$f_R(\|\mathbf{y}\|) = f_{\mathbf{Y}}(\mathbf{y}) S_{N-1}(\|\mathbf{y}\|), \quad (9)$$

where

$$S_{N-1}(r) = \frac{2\pi^{N/2} r^{N-1}}{\Gamma(N/2)} \quad (10)$$

is the area of the hypersphere  $P_r$  [16, Sec. V].

An integral over  $\mathbb{R}^N$  can be written as a double integral: an outer integral over  $r$  from 0 to  $\infty$  and an inner integral over  $P_r$ . Therefore, (3) can be written as

$$h(\mathbf{Y}) = - \int_0^\infty \int_{P_r} f_{\mathbf{Y}}(\mathbf{y}) \log_2 f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} dr \quad (11)$$

$$= - \int_0^\infty \int_{P_r} \frac{f_R(\|\mathbf{y}\|)}{S_{N-1}(\|\mathbf{y}\|)} \log_2 \frac{f_R(\|\mathbf{y}\|)}{S_{N-1}(\|\mathbf{y}\|)} d\mathbf{y} dr \quad (12)$$

$$= - \int_0^\infty \frac{f_R(r)}{S_{N-1}(r)} \log_2 \frac{f_R(r)}{S_{N-1}(r)} dr \int_{P_r} d\mathbf{y} \quad (13)$$

$$= - \int_0^\infty f_R(r) \log_2 \frac{f_R(r)}{S_{N-1}(r)} dr \quad (14)$$

$$= h(R) + \mathbb{E}[\log_2 S_{N-1}(R)], \quad (15)$$

where (12) follows from (9).<sup>2</sup> Substituting (10) in (15) completes the proof of (6).

Let now  $T = R^N$ . By standard rules for transformation of random variables,  $f_R(r) = N r^{N-1} f_T(r^N)$  for all  $r \geq 0$ . Taking logarithms,

$$\log_2 f_R(R) = \log_2 f_T(T) + \log_2 N + (N-1) \log_2 R. \quad (16)$$

Negating and taking expectations yields

$$h(R) = h(T) - \log_2 N - (N-1) \mathbb{E}[\log_2 R]. \quad (17)$$

Substituting (17) in (6) and using the gamma function identity  $u\Gamma(u) = \Gamma(u+1)$  proves (7).  $\square$

Using Lemma 1, an expression for the mutual information between the input and output of the AWGN channel with a given rotationally invariant input distribution can be derived as follows.

**Theorem 2:** If  $\mathbf{X}$  has a rotationally invariant distribution, characterized by the distribution  $f_{\|\mathbf{X}\|}$ , then the mutual information between  $\mathbf{X}$  and  $\mathbf{Y}$  in (1) in bits per  $N$ -D channel use is

$$I(\mathbf{X}; \mathbf{Y}) = - \int_0^\infty f_{\tilde{R}}(\tilde{r}) \log_2 \frac{f_{\tilde{R}}(\tilde{r})}{\tilde{r}^{N-1}} d\tilde{r} + \log_2 \frac{2}{\Gamma(N/2)} - \frac{N}{2} \log_2 2e, \quad (18)$$

<sup>2</sup>In [17, Eq. (4.12)], an expression similar to (14) was proved for a unit hypersphere. However, the proof above appears to be simpler.

where

$$f_{\tilde{R}}(\tilde{r}) = \int_0^\infty f_{\|\mathbf{X}\|}(s) \chi\left(\tilde{r}, \frac{s}{\sigma}\right) ds, \quad (19)$$

$$\chi(\tilde{r}, \tilde{s}) = \frac{\tilde{r}^{N/2}}{\tilde{s}^{N/2-1}} \exp\left(-\frac{\tilde{r}^2 + \tilde{s}^2}{2}\right) I_{N/2-1}(\tilde{r}\tilde{s}), \quad (20)$$

and  $I_\nu(u)$  is the  $\nu$ th order modified Bessel function of the first kind.

*Proof:* Substituting (4) and (6) in (2) and using  $R = \|\mathbf{Y}\|$  yields

$$I(\mathbf{X}; \mathbf{Y}) = - \int_0^\infty f_R(r) \log_2 \frac{f_R(r)}{r^{N-1}} dr + \log_2 \frac{2\pi^{N/2}}{\Gamma(N/2)} - \frac{N}{2} \log_2 2\pi e \sigma^2 \quad (21)$$

$$= - \int_0^\infty f_R(r) \log_2 \frac{\sigma^N f_R(r)}{r^{N-1}} dr + \log_2 \frac{2}{\Gamma(N/2)} - \frac{N}{2} \log_2 2e. \quad (22)$$

Defining  $\tilde{R} = R/\sigma$ , we substitute  $f_R(r) = (1/\sigma) f_{\tilde{R}}(r/\sigma)$  and  $r = \sigma \tilde{r}$  in (22) to obtain (18).

From [18, Eqs. (1.14), (2.44), and (2.48)], the conditional distribution of  $Q = \|\mathbf{Y}\|^2$  given  $\|\mathbf{X}\| = s \geq 0$  is the noncentral chi-square distribution with  $N$  degrees of freedom

$$f_{Q|\|\mathbf{X}\|}(q|s) = \frac{1}{2\sigma^2} \left(\frac{q}{s^2}\right)^{N/4-1/2} \exp\left(-\frac{q+s^2}{2\sigma^2}\right) \cdot I_{N/2-1}\left(\frac{s\sqrt{q}}{\sigma^2}\right), \quad q \geq 0. \quad (23)$$

Substituting  $Q = \sigma^2 \tilde{R}^2$  yields the noncentral chi distribution

$$f_{\tilde{R}|\|\mathbf{X}\|}(\tilde{r}|s) = 2\sigma^2 \tilde{r} f_{Q|\|\mathbf{X}\|}(\sigma^2 \tilde{r}^2 | s) = \chi\left(\tilde{r}, \frac{s}{\sigma}\right), \quad \tilde{r} \geq 0, \quad (24)$$

with  $\chi$  defined in (20). Finally, (19) follows by  $f_{\tilde{R},\|\mathbf{X}\|}(\tilde{r}, s) = f_{\|\mathbf{X}\|}(s) f_{\tilde{R}|\|\mathbf{X}\|}(\tilde{r}|s)$  using (24).  $\square$

The following corollaries confine Theorem 2 to the special case of multisphere input distributions. The integral in (19) is replaced by sums, which reduces the computational complexity.

**Corollary 3:** Let  $\mathbf{X}$  be distributed according to an  $N$ -D multisphere distribution, where the probabilities and radii of each hypersphere  $k = 1, \dots, K$  are  $p_k$  and  $s_k$ , respectively. Then the mutual information between  $\mathbf{X}$  and  $\mathbf{Y}$  in bits per  $N$ -D channel use is given by (18), (20), and

$$f_{\tilde{R}}(\tilde{r}) = \sum_{k=1}^K p_k \chi\left(\tilde{r}, \frac{s_k}{\sigma}\right). \quad (25)$$

*Proof:* Trivial from Theorem 2.  $\square$

**Corollary 4:** Let  $\mathbf{X}$  be distributed according to an  $N$ -D multisphere distribution with  $K$  uniformly spaced hyperspheres and equal probabilities per hypersphere. Then the mutual information between  $\mathbf{X}$  and  $\mathbf{Y}$  in bits per  $N$ -D channel

use as a function of the SNR  $A = \mathbb{E}[\|\mathbf{X}\|^2]/N_0$  is given by (18), (20), and

$$f_{\tilde{R}}(\tilde{r}) = \frac{1}{K} \sum_{k=1}^K \chi(\tilde{r}, k\Delta), \quad (26)$$

$$\Delta = \sqrt{\frac{12A}{2K^2 + 3K + 1}}. \quad (27)$$

*Proof:* Let  $p_k = 1/K$  and  $s_k = k\sigma\Delta$  for  $k = 1, \dots, K$  in Corollary 3. The hypersphere spacing  $\Delta$  is determined from

$$\mathbb{E}[\|\mathbf{X}\|^2] = \sum_{k=1}^K p_k s_k^2 \quad (28)$$

$$= \frac{\sigma^2 \Delta^2}{K} \sum_{k=1}^K k^2 \quad (29)$$

$$= \frac{\sigma^2 \Delta^2 (2K^2 + 3K + 1)}{6}. \quad (30)$$

Equating (30) with  $\mathbb{E}[\|\mathbf{X}\|^2] = 2\sigma^2 A$  yields (27).  $\square$

As a further special case, setting  $K = 1$  and  $N = 2$  in Corollary 3 or 4 gives the constrained capacity of constant-amplitude constellations, i.e., purely phase-modulated transmission schemes, in the presence of AWGN, which was derived by Wyner [19, Eq. (12)].

At high SNR, the integrand in (18) is very peaky for multisphere input distributions, which can cause numerical problems. These problems can be effectively avoided by sampling these peaks, whose locations are  $\tilde{r} \approx \Delta, 2\Delta, \dots, K\Delta$ , denser in the integration algorithm.

#### IV. EVALUATION USING 2- AND 4-D MULTISPHERE DISTRIBUTIONS

We evaluate Corollary 4 by numerical integration for a distribution of  $K$   $N$ -D hyperspheres, where  $K = \{1, 2, 4, 8\}$ ,  $N = \{2, 4\}$ , and the probability of transmitting a point belonging to a certain hypersphere is uniform. The specific choices of  $N$  are motivated by coherent fiber-optic communication with its 4-D signal space, and noise that can be well approximated by AWGN [20]. We compare a 2-D multisphere distribution transmitted on both polarizations, i.e., using the 4 available dimensions, and a 4-D multisphere distribution.

The results are shown in Fig. 2, along with the unconstrained capacity (5) of an  $N$ -D AWGN channel with  $N = 2$  and 4. At low to medium SNR, the multisphere distributions perform quite close to capacity even with only 2 or 4 hyperspheres. The behavior at high SNR, however, differs. It can be observed in Fig. 2 (a) that the mutual information of the 4-D and 2-D multisphere distribution asymptotically grows as  $(3/2) \log_2 A$  and  $(1/2) \log_2 A$ , respectively, whereas the corresponding capacities (5) grow as  $2 \log_2 A$  and  $\log_2 A$ , respectively. The coefficient of  $\log_2 A$ , called the *prelog*, is, in analogy with the well-known expression for the multiplexing gain in multiple-input multiple-output communications, equal to the number of complex *degrees of freedom* available for transmission [21, Chs. 7–9]. The  $N$ -D unconstrained AWGN channel (1) has  $N$  real degrees of freedom, or equivalently  $N/2$  complex degrees of freedom, which explains the observed

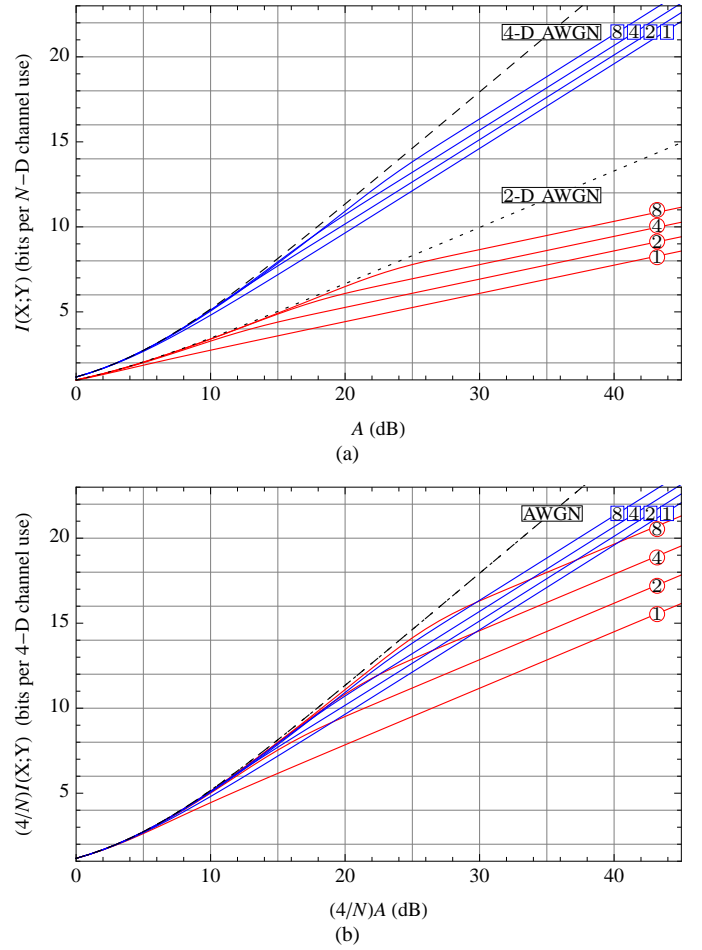


Fig. 2. (a) Achievable rates in bits per channel use of  $N$ -D multisphere distributions, for  $N = 2$  (circle) and 4 (rectangle), plotted as functions of SNR  $A$ . (b) The same curves normalized to show rates in bits per 4-D channel use as functions of 4-D SNR (i.e., the ratio between energy per four dimensions and  $N_0$ ).

prelogs 2 and 1 for  $N = 4$  and 2, respectively. On the other hand, an  $N$ -D multisphere distribution has effectively only  $N - 1$  real degrees of freedom, because the input vector can vary continuously along all  $N$  dimensions except the amplitude, which is discrete; this yields a prelog of  $(N - 1)/2$ , which again agrees with the observations in Fig. 2 (a).

A perhaps more relevant scenario is to assume that 4 dimensions are used in all cases, with a fixed amount of energy. The input distribution is either a single 4-D multisphere distribution or two multiring distributions used in parallel, with half the energy each. Such a comparison is shown in Fig. 2 (b), where both axes in Fig. 2 (a) have been normalized by a factor  $4/N$ . In this scenario, the two Gaussian results coincide, since a 4-D Gaussian distribution is the same as two independent 2-D Gaussian distributions. At high SNR, 4-D multisphere distributions are clearly superior to their 2-D counterparts. This can again be explained by their prelogs, which govern the high-SNR slopes. A single 4-D multisphere distribution has 3 real degrees of freedom, whereas two 2-D multiring distributions have only 2 real degrees of freedom together; hence the asymptotic prelogs are  $3/2$  and 1, respectively, whereas the 4-D Gaussian distribution

still has a prelog of 2. Interestingly, the gain from increasing the number of hyperspheres  $K$  is higher if  $N = 2$  than 4. This has the somewhat counterintuitive effect that in some cases, for example at 25 dB and  $K = 8$ , two independent 2-D multisphere distributions have a better performance than a single 4-D multisphere distribution. The reason is that the combination of two 2-D distributions yields an amplitude distribution that is more similar to that of a Gaussian 4-D distribution.

## V. STATISTICAL ROTATIONAL INVARIANCE OF THE STOCHASTIC MANAKOV EQUATIONS

In addition to AWGN resulting from optical amplifiers, transmission over optical fibers is impacted by fiber nonlinearity. Therefore, it is of great interest to investigate how 4-D multisphere distributions are impacted by such nonlinearity. In this section, we show an interesting *statistical* invariance of each point on a 4-D sphere of a multisphere distribution by showing the statistical invariance of the *nonlinear* propagation equations of the two complex fields present in optical fibers.

A good approximation for the equations of propagation of the two polarization states supported in a single-mode fiber is a set of coupled partial differential equations referred to as *Manakov equations* [22]. They are generally written in a vector form [23],

$$\frac{\partial \mathbf{E}}{\partial z} + \imath \frac{\beta_2}{2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \imath \gamma \frac{8}{9} \|\mathbf{E}\|^2 \mathbf{E} = \imath \mathbf{N}, \quad (31)$$

where  $\mathbf{E}(z, t) = [E_x(z, t) \ E_y(z, t)]^T$  with  $E_x(z, t)$  and  $E_y(z, t)$  being the two complex fields in the  $x$  and  $y$  polarizations, respectively,<sup>3</sup> as functions of distance  $z$  and time  $t$ , and where  $\mathbf{N}(z, t) = [N_x(z, t) \ N_y(z, t)]^T$  with  $N_x(z, t)$  and  $N_y(z, t)$  being the  $x$  and  $y$  components of the noise. The noise field  $\mathbf{N}(z, t)$  is white in  $z$  and  $t$  and has a circular Gaussian distribution in each quadrature.

Consider the transformation

$$\mathbf{E}(z, t) = \mathbf{U} \mathbf{E}'(z, t), \quad (32)$$

where  $\mathbf{U}$  is an arbitrary 2-D (complex) unitary matrix that is independent of  $z$  and  $t$ . This transformation preserves the norm of  $\mathbf{E}(z, t)$ , i.e.,  $\|\mathbf{E}'\| = \|\mathbf{E}\| = \sqrt{E_{x,r}^2 + E_{x,i}^2 + E_{y,r}^2 + E_{y,i}^2}$ , where the indices  $r$  and  $i$  denote the real and imaginary components of the fields, respectively. Preserving the norm means that all points on a hypersphere in the 4-D space of real fields remain on that hypersphere. Substituting (32) in (31) and multiplying the equation by  $\mathbf{U}^{-1}$ , the new equation expressed in  $\mathbf{E}'(z, t)$  assumes the same form as (31) except for the noise term  $\mathbf{N}(z, t)$  on the right-hand side, which is now  $\mathbf{U}^{-1} \mathbf{N}(z, t)$ . One can show that any 2-D unitary transform operating on a 2-D complex AWGN field  $\mathbf{N}(z, t)$  preserves its statistical properties [24, Sec. 7.8.1]. As a result, all 2-D unitary transformations of the 2-D complex vector field  $\mathbf{E}(z, t)$  leave the Manakov equations (31) statistically invariant. Consequently, all points lying on the same hypersphere of a 4-D multisphere constellation can be considered as statistically equivalent, even in the presence of fiber nonlinearity. One can exploit this

equivalence, for example, to reduce the number of data realizations (i.e. symbols) needed to numerically compute a capacity estimate of a nonlinear system governed by (31) [25]. One can show that the statistical invariance demonstrated in this section can be readily extended to include additional linear terms in the field  $\mathbf{E}(z, t)$  in (31), including higher-order dispersions, fiber loss and amplifier gain. It is interesting to note that a spherical constellation of the form of a multidimensional ball (no discretization of radii) has been studied and shown to provide some level of mitigation of nonlinear effects [26].

## VI. CONCLUSIONS

The mutual information of  $N$ -D rotationally invariant input distributions is derived in the presence of additive white Gaussian noise, by exploiting the rotationally invariant properties of the noise and the distribution itself. Targeting applications in coherent fiber-optic transmission systems dominated by linear noise, where the signal space is 4-D due to the two quadratures in both polarizations of the optical carrier, the input distribution is confined to a finite number of concentric, uniformly spaced, equally probable  $N$ -D hyperspheres. At high SNR, it is shown that a 4-D multisphere distribution enables higher achievable rates than a 2-D multisphere distribution transmitted on both polarizations, even though both distributions make use of the available 4-D signal space. Finally, the statistical invariance of the 4-D multisphere distribution was shown under the set of partial differential equations describing the nonlinear evolution of the optical fields in optical fibers.

## VII. ACKNOWLEDGMENTS

We would like to thank G. J. Foschini for his enlightening comments that helped provide a proper direction of this work. We would also like to thank R. W. Tkach for his continuous support.

## REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, 1948.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey: John Wiley & Sons, Inc., 2006.
- [3] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1060–1071, July 1995.
- [4] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inf. Control*, vol. 18, pp. 203–219, 1971.
- [5] M. Katz and S. Shamai (Shitz), "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.
- [6] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.
- [7] M. C. Gursoy, H. V. Poor, and S. Verdú, "The noncoherent Rician fading channel—Part I: Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2193–2206, Sept. 2005.
- [8] H. Ghazlan and G. Kramer, "Interference focusing for mitigating cross-phase modulation in a simplified optical fiber model," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, June 2010, pp. 2033–2037.
- [9] —, "Interference focusing for simplified optical fiber models with dispersion," in *Proc. IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, July–Aug. 2011, pp. 376–379.

<sup>3</sup>We use  $\|\cdot\|$  in (31) for consistency with the rest of this paper even though this equation is more commonly written using  $|\cdot|$  in the physics community.

- [10] T. Freckmann, R.-J. Essiambre, P. J. Winzer, G. J. Foschini, and G. Kramer, "Fiber capacity limits with optimized ring constellations," *IEEE Photon. Technol. Lett.*, vol. 21, no. 20, pp. 1496–1498, Oct. 2009.
- [11] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.
- [12] M. I. Yousefi and F. R. Kschischang, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.
- [13] E. Agrell, "Nonlinear fiber capacity," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, London, UK, Sept. 2013.
- [14] A. Alvarado and E. Agrell, "Four-dimensional coded modulation with bit-wise decoders for future optical communications," *J. Lightw. Technol.*, vol. 33, no. 10, pp. 1993–2003, May 2015.
- [15] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
- [16] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, May 1959.
- [17] G. J. Foschini and A. H. Diaz, "Dirty paper coding: perturbing off the infinite dimensional lattice limit," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 62, pp. 141–160, 2003.
- [18] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. New York, NY: Springer, 2002.
- [19] A. D. Wyner, "Bounds on communication with polyphase coding," *Bell Syst. Tech. J.*, pp. 523–559, Apr. 1966.
- [20] J. P. Gordon, W. H. Louisell, and L. R. Walker, "Quantum fluctuations and noise in parametric processes II," *Phys. Rev.*, vol. 129, no. 1, pp. 481–485, 1963.
- [21] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge University Press, 2005.
- [22] S. V. Manakov, "Contribution to the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Sov. Phys. JETP*, vol. 38, no. 2, pp. 248–253, 1974.
- [23] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, 1997.
- [24] R. G. Gallager, *Principles of Digital Communication*. Cambridge, U.K.: Cambridge University Press, 2008.
- [25] R.-J. Essiambre, R. W. Tkach, and R. Ryf, "Fiber nonlinearity and capacity: Single-mode and multimode fibers," in *Optical Fiber Telecommunications VI B*, I. Kaminow, T. Li, and A. E. Willner, Eds. Academic Press, 2013, ch. 1, pp. 1–37.
- [26] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," *Proc. IEEE Int. Symp. Inf. Theory*, pp. 2794–2798, 2014.